

NVIDIA Tesla P100 GPU 助力冷冻电镜云计算平台

案例简介

- 本案例中，上海承蓝科技股份有限公司利用NVIDIA GPU助力冷冻电镜云计算平台系统的搭建：首先利用Tesla P100 GPU的计算极大地加速了科研数据分析的计算速度，对加速解析生物大分子的三维结构起着至关重要的作用；其次利用电镜云计算分布式架构技术构建计算数据共享组的概念，高效利用GPU和高性能存储资源。

- 本案例主要应用到拥有NVIDIA Tesla P100 GPU 的超级计算机。

Case Introduction

- In this case, Shanghai Chenglan Technology Co.,Ltd applies the NVIDIA GPU to build the Cryo-EM Cloud Platform. Firstly, the Tesla P100 dramatically accelerate the speed of research data analysis, and it plays the key point in the analysis process of biomacromolecule three-dimensional structures. Secondly, the Cryo-EM Cloud Platform implemented by distributed and multi-tier method builds the concept of computing data sharing group, which can efficiently utilize the GPU and high-performance storage resources.

- The major product utilized in the case is supercomputer with 4 NVIDIA Tesla P100 GPU.

现状

上海承蓝科技股份有限公司（以下简称“承蓝科技”），成立于2005年3月，2016年8月在“新三板”挂牌上市。承蓝科技是一家专注于生命科学与医疗大数据平台的高新技术企业，致力于融合数据挖掘、深度学习、人工智能等前沿科技为医疗与生命科学领域基础研究、疾病智能辅助诊断等领域提供全方位的工具和服务。秉承“唤醒数据”的使命及“为用户创造价值”的核心理念，承蓝科技始终为广大科研、医疗用户提供更加简单、高效、智能、安全的服务，让数据产生价值。

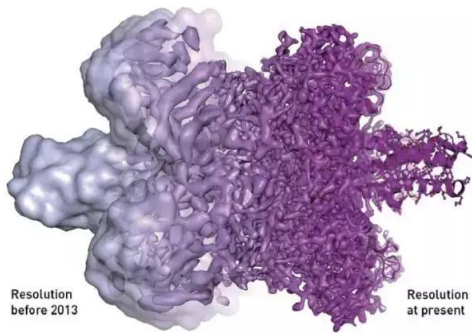
冷冻电镜方法早在上世纪六十年代就开始发展，但是由于技术方法的瓶颈，一直只能做一些较低分辨率的结构解析工作。近年来，冷冻电镜技术飞速发展，不断突破，在分辨率上已经开始可以和晶体学相媲美。而且，由于其不需要结晶，对样品的均一性要求也相对较低，样品用量少，可重复性较高，加上快速冷冻能使生物分子尽量保持其天然结构状态，冷冻电镜三维重构技术已经成为结构生物学领域一颗亮眼的明星。

冷冻电镜方法耗时最长的环节是数据的分析与处理，可能要花费一个专业人员数周以上的的时间。对于一个包含数百万个颗粒，拥有多个构像分子量较大的数据，即使在高性能计算集群上，也可能要花费超过50万CPU小时的时间。引入GPU加速技术是目前很多主流软件的选择，在GPU的加速帮助下，大大缩短了分析计算的时间与成本。

承蓝科技的冷冻电镜云平台的方案计算单元基于NVIDIA Tesla GPU提供的卓越计算性能，将其应用于冷冻电镜数据的三维结构解析，全面提升数据计算的速度和结构构象的分辨率。

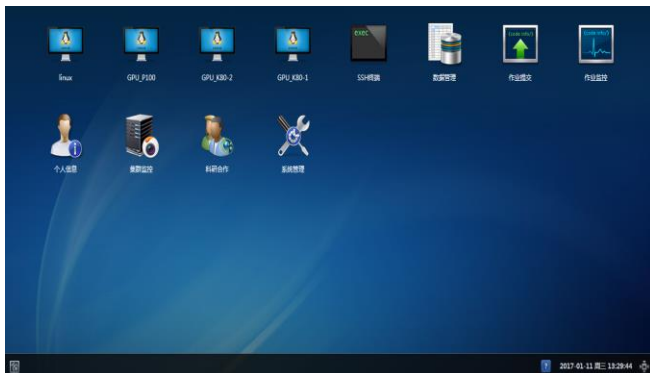
挑战

从电子显微镜的发明，到冷冻电镜技术在80年代的基本确立，经历了将近50年的时间，再到原子分辨率三维重构的普遍实现，又经历了将近40年的时间。这将近90年的发展过程从一个侧面体现了科学家们在发展冷冻电镜技术的过程中所面临的巨大挑战。正是由于一大批科学家和技术人员的不懈努力，才成就了冷冻电镜这一个划时代的技术。从三维重构原理的提出，到样品快速冷冻技术确立，再到图像分析技术的逐步成熟完善，生物样品的辐照损伤问题被逐渐地解决，最终使冷冻电镜技术逐渐成为一个具有巨大潜力的革命性技术。这一过程与包括生物学、物理学、数学、电子工程和计算机技术等在内的多学科交叉是密不可分的。随着新方法新技术的不断融入，或许在不远的将来，科学家们将可以利用冷冻电镜在细胞中直接测定生物大分子的原子结构，揭示隐藏在结构之中的生命奥秘。



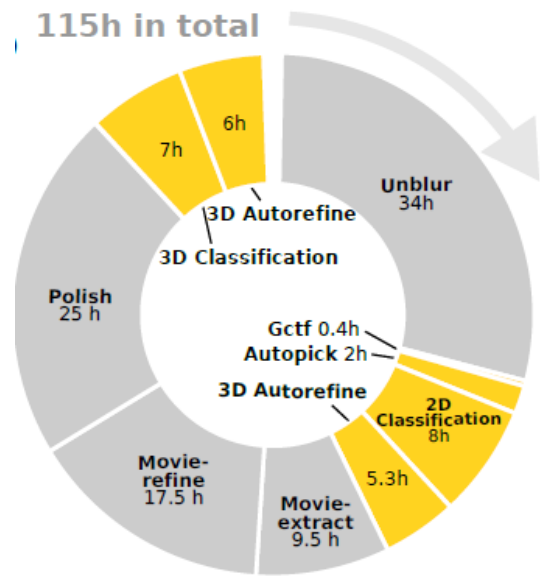
整体方案

承蓝电镜云计算平台结合 Nvidia Tesla P100 GPU 提供冷冻电镜数据规模化分析、海量数据生命周期等功能模组。支持用户在线利用高性能 GPU 节点完成数据解析数据录入、分析与结果展示。支持跨单位、跨课题组、跨用户间的科研合作项目数据安全、协议管理、人员管理及审批管理，提供用户身份、数据管理与文件系统统一整合。本方案已经成功部署在复旦大学、北京大学等冷冻电镜平台。



NVIDIA Tesla P100 GPU 加速器为现代数据中心释放强大的计算能力。它利用全新的 NVIDIA Pascal™ 架构打造出速度极快的计算节点，性能高于数百个速度较慢的通用计算节点。利用更少的快速的节点获得更高的性能，能在节省资金的同时，大幅提高数据中心吞吐量。

NVIDIA Tesla P100 GPU 为冷冻电镜数据分析提供了强大且稳定的计算能力。基于 P100 的三维重构，相比于基于 CPU 的计算，有平均大于 15 倍的提升。这使在较短时间内可以得到三维结构信息以及其变化过程中的可能的构象，这为用于下一步的科研提供了坚实的基础。



影响

2015 年 Nature Methods 将冷冻电镜技术评为年度最受关注的技术，2017 年冷冻电镜技术获得诺贝尔化学奖实至名归。今后，冷冻电镜单颗粒技术的应用必然会更加普及，对于重要生物大分子的结构解析课题组之间的竞争将更加白热化。

对冷冻电镜三维重构技术，受益于 P100 GPU 的计算能力，得以在短时间内展开了新的思路。通过 GPU 并行加速，实现蛋白颗粒的分组，结构重构以及结果优化。